IMPROVING THE PERFORMANCE OF DELIBERATIVE GROUPS THROUGH CHANGES IN ORGANIZATIONAL STRUCTURE: A SIGNAL DETECTION APPROACH

By

JESSE ITZKOWITZ

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2005

This dissertation is dedicated to my wife, Jennifer, the love of my life and my best friend.

ACKNOWLEDGMENTS

I would like to thank the members of my doctoral committee for their patience, support, guidance, and understanding. Foremost, I would like to thank Dr. Robert D. Sorkin. There is no doubt that Dr. Sorkin has made me a better scientist. Throughout the doctoral process, Dr. Sorkin provided me with scholarly wisdom, an independent work ethic, and countless ideas for future research. I would also like to thank Dr. Ira Fischler, who, aside from first introducing me to the field of cognitive psychology, served as my cochair during Dr. Sorkin's sabbatical. I really appreciated his open door policy and helpful guidance. Special thanks go to Dr. Joseph Alba and to Dr. Mark Fondacaro. Both had constant enthusiasm for my work which served as great encouragement throughout my doctoral studies. I appreciated their ability to help me bridge the gap between my theoretical work to broader legal and business issues.

I would also like to thank Jennifer Itzkowitz, my wife. There is no way that I could possibly thank her for all she has done to help me through the doctoral process. Helping test my experimental programs went well beyond the traditional call of duty for any spouse. Her confidence in me and her support for me were unparalleled. I would also like to thank Jennifer for her comments, ideas, and suggestions on the dissertation.

Finally, I thank my parents, Gary and Doris Itzkowitz. They have always encouraged me to seek out my interests, wherever they take me. Throughout my doctoral education here, my parents were always willing to listen to me talk about my work, even

when I am positive that they were unsure exactly what I was talking about. Their unwavering support for me has been an essential component of my success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

IMPROVING THE PERFORMANCE OF DELIBERATIVE GROUPS THROUGH CHANGES IN ORGANIZATIONAL STRUCTURE: A SIGNAL DETECTION APPROACH
By

Jesse Itzkowitz

August 2005

Chair: Robert D. Sorkin
Cochair: Ira Fischler
Major Department: Psychology

Group deliberation, whereby members share their individual judgments of a binary event to each other one at a time, while simultaneously updating their own opinions accordingly, is commonplace in group decision making. Examples of this type of deliberation can be seen in juries, business meetings, and military teams. We present a Signal Detection Model of the deliberation process. This model posits that group deliberation is optimized when individual members recalculate their individual decision criterion in a Bayesian manner based upon the source and value of the current opinion provided to the group. Consistent with our model, results of actual groups indicate that group members update their own opinions similar to the prescriptions of the model.

Additionally, we investigated the interaction between groups' organizational structure and groups' consensus accuracy. Specifically, we examined groups in which deliberation was sequenced a) by member response time, b) by member ability, c) by a

proxy for member confidence, and d) at random selection. We also investigated the impact of both a majority and unanimous requirement for group consensus. Results indicated that measures which caused the group to deliberate for longer amounts of time led to correspondingly higher group accuracy.

CHAPTER 1
INTRODUCTION

Consider the following scenarios: you are asked to determine whether an individual is innocent or guilty of an alleged crime; you must determine whether or not to invest your company's money into a risky proposal; or you must decide if there is evidence to take preemptive military action against a nation that may be planning a military strike of its own. For all of these scenarios, would you rather make this decision independently or make this decision in a group?

Groups have often been called upon to make important decisions due to uncertain individual judgment, due to the magnitude of decisional impact, or due to the complexity of the decision itself for two main reasons. First, groups allow individuals to distribute responsibility among its members so that no one member bears the burden (or basks in the glow) of any consequential group decision (Gigone & Hastie, 1996). Thus group decision making allows individuals to feel less responsible for incorrect decisions. Secondly, groups are chosen to make difficult decisions because they are believed to possess greater variation of opinions, abilities, and biases than individuals alone, thus leading to more accurate decisions. This variation between group members allows groups to make better decisions than could be made by any one individual when members are asked to combine each other's private, independent information with one another before reaching decision (Stasser & Titus, 1985).

One decision environment that follows this procedure is group deliberation. We define deliberation as the process by which group members receive private information

about an event and are then asked to make their opinions public, through sequential, serial disclosure. Deliberative groups as described here are quite commonplace. The deliberative process is often used in juries, military command and control systems, and in medical teams. These deliberative groups are not bound by physical location. While members can meet face-to face, members can also be distributed, as in financial markets (where buying and selling represent the group members' opinions) and distributed detection systems (such as those used in radio astronomy or weather prediction).

Because of the sequential nature of deliberation, there are often rules imposed that govern its course. We label the set of rules that specify the group's specific speaking orders as the group's *response protocol* and rules that define consensus (and thus end the deliberative process) as the group's *decision criteria*. Taken together, a group's response protocol and decision criteria compose its *organizational structure*.

In this study, we hope to answer three essential questions about groups. First, how should groups decide? Second, how do groups actually decide? Third, what prescriptions can be made to improve the performance of actual groups? Using Signal Detection Theory (SDT), we have modeled the process of deliberation, which defines how group members should use the information provided to them from other group members. Following the presentation of the model, we investigate groups with natural deliberation, and discover how groups behave without more invasive organizational structures (e.g., when deliberation sequences reflect member desire to respond). Finally, using the SDT model, we simulate group behavior using some of the most ubiquitous organizational structures. We show that these simulations provide reasonably good interpretations of actual group performance, demonstrating that group members perform

deliberation in a rational manner. Finally, we discuss interventions that should increase the performance of deliberative groups.

Before we present the SDT model, it is important to consider the effect of deliberation. We begin by examining previous attempts at modeling the group deliberation process. Following this discussion, we move to discuss techniques for aggregating information in groups. Indeed, there may be any number of ways that member information can be combined within the group in order to reach a decision. We examine these various techniques and discuss implications for normative modeling.

Models of Group Deliberation

Although the group deliberation process is highly complex and dynamic, there have been numerous attempts to model both the process and decision outcome of deliberative groups (Davis, 1973; Davis, Stasson, Ono, & Zimmerman, 1988; Davis, Stasson, Parks, Hulbert, Kameda, Zimmerman, & Ono, 1993; Davis, Tindale, Nagao, Hinsz, & Robertson, 1984; Kerr, MacCoun, & Kramer, 1996; Penrod & Hastie, 1980). These studies have allowed the investigation of a number of different phenomena that govern group behavior such as group size, decision rule, and the dynamics of majority and minority influence within groups.

While all of these studies differ in the types of hypotheses tested and the results found, all use variations of the Social Decision Scheme (SDS) theory developed by Davis (1973). The SDS theory of group decision making describes how the initial distribution of members' opinions within the group interacts to form the final group decision. Here, decisional outcomes are modeled based on decision schemes which aggregate members' initially stated opinions into final group judgments. These decision schemes are

essentially different group decision criteria. Probabilistic distributions of initial member opinion serve as the input functions to these schemes.

While none of the current SDS examinations parallel our definition of deliberation, three provide insight into the effects of organizational structure and group composition. Davis et al. (1998) examined the influence of sequential voting in groups. In this study, groups of six were divided between members that preferred guilty or not guilty verdicts in a mock trial. Groups performed a sequential voting procedure that either began through disclosure of three guilty votes or three not-guilty votes, depending on condition. It was found that these votes had a significant effect on both group judgment and the individual opinions of the group members who voted after presentation of the other jurors' opinions. Explanations for this phenomenon rested upon the influence of majority factions within the group. Members who voted subsequent to the initial presentation of votes were believed to be participating in a conforming process to acquiesce to the majority faction. Thus, sequential presentation of votes within decision making groups has a substantial impact on the opinions of the group's members.

In a similar study of juries, Davis et al. (1993) investigated how sequence of votes affected performance in a continuous judgment task (damage awards in a mock civil trial). Again, it was shown that the sequence of presentation of member opinion influenced the final damage award decisions made by the juries. Most importantly, it was found that the first votes presented to the jury had substantially more influence than later votes (as determined through examination of polls taken early and late in the deliberative process).

Finally, Kerr, MacCoun, and Kramer (1996) use the SDS method to provide the most comprehensive examination of how group composition, group bias, and decision processes interact in final judgment. Results determined that final group bias was determined by the type of bias and degree of bias the individual group members possessed, along with the type of decision scheme implemented. While attempts to determine whether or not groups had larger or smaller bias than their members were mixed, it was found that the initial biases of the group members had significant impact on the groups' final decision, regardless of the decision scheme used. Thus any normative account of group decision making must incorporate individual member bias in some way.

While these SDS models allow us to understand some of the dynamics of the deliberative process through these examinations, there are two disadvantages to the methodology which we hope to eliminate in our Signal Detection Model. First, SDS models are not integrated models of individual and group judgment. In all cases, it is assumed that the individual group members have already reached some sort of decision before the model can be applied. The Signal Detection methodology allows us to incorporate the process of individual judgment into the group decision process. This ensures that information about member bias and expertise is inherent in the model. Second, SDS models do not usually allow for the accuracy of groups to be determined (Hastie & Kameda, 2005). While the models are useful in allowing us to determine the proportion of times a group reaches a particular decision, they do not allow us to measure the effectiveness of any given decision scheme. In contrast, because the SDT model is based on a basic detection task, measurements of accuracy are an essential part of the system. In conclusion, SDS models are a useful way of determining group decision if the

choices of the group's members are known and if measures of accuracy are not needed. Although these models answer some of the basic questions regarding the influence of groups' organizational structures on their final decision performance, they do not lead us to methodologies of combining group opinion in a normative manner.

Combining Information in Groups

Within group decision making, different techniques exist for aggregating group decision and reaching consensus. These methods vary widely in performance and structure based on the type of task (including type of information), and the goals of the group (Clemen, 1989; Clemen & Winkler, 1999; Papastavrou & Athans, 1992). Before determining which type of aggregation method should be used by the group, it is critical to determine if the group task is either intellective or judgmental. A task is intellective to the extent to which it relies on information and that there are rules to evaluate the correctness of a certain answer (Kaplan & Martin, 1999; Zarnoth & Sniezek, 1997). Judgmental tasks have no demonstrable correct answer and are often thought of as matters of opinion (Hinsz, Tindale, & Vollrath, 1997; Zarnoth & Sniezek 1997). Because we are interested in discussing the accuracy of group performance, we will focus on those techniques that combine member opinion for intellective tasks.

Aside from determining which type of task the group is performing, it is also important to determine the group's goals. For most groups, this goal is accuracy (Einhorn, Hogarth, & Klempner, 1977; Grofman, Owen, & Feld, 1983; Sorkin & Dai, 1994; Watson, Michaelson, & Sharp, 1991). However, for other groups, this goal could be fairness, autonomy, speed, or ease of decision making. Again, because we are interested in how groups should operate in order to reach the most correct decisions, we will focus only on techniques for combining opinion that improve decisional accuracy.

The literature provides two main methods for combining member opinion within groups: behavioral and mathematical. Behavioral aggregation techniques attempt to generate agreement among the experts by having them interact (Clemen & Winkler, 1999). Behavioral methods of information combination generally assume that information can be aggregated and conflicts of opinion can be settled simply through group discussion. However, as research has shown, discussion can only be useful in shifting other members' opinions to the extent that it is a reasonable and logical exchange of information; else decision makers will maintain their initial positions (Innami, 1994). However, when there is no clear structure to the informational discussion inherent in intellective group tasks, groups are more likely to fall prey to normative effects of group pressure because there is no clear structure to the aggregation technique, thus decreasing group performance.

Because we are interested in a normative solution to the group combination problem, we gain more insight from mathematical methods of combining member opinion. These techniques use the information provided by the group's members along with information about the group member's decisional characteristics to combine opinions according to some sort of mathematical function. These mathematical functions can take on many forms, and there have been quite a few techniques that describe mathematical aggregation within groups (e.g., Grofman & Feld, 1986; Hoballah & Varshney, 1989; Nitzan & Paroush, 1982; Pete, Pattipati, & Kleinman, 1993). In the majority of these techniques, the individual decisions made by the group members are aggregated outside the further influence of the members and group members may not have approval over the final decision. The site of aggregation has different names

according to discipline such as fusion center, oracle, judge, or head decision maker. Here, according to some prescribed method of integration and decision, a consensus decision is made. All of the reviewed models weight the decisions of the more expert group members more heavily than the decisions of the less expert group members, but Pete et al. (1993) show that the optimal combination of member opinion comes though a Bayesian combination of the members' responses (likelihood ratios) incorporating information about other members' expertise and bias. While this finding is useful, in that it provides a normative standard of how to weight the tally of the group's votes, it assumes a one-vote, parallel process rather than the process of deliberation described earlier. Additionally, our definition of deliberation requires that the decision is made by the group members thus any technique that calls for an outside "fusion center" does not meet our strict definition of group deliberation.

Interestingly, while studying the effects of Bayesian recalculation at the fusion center, it was discovered that optimal group performance would come from Bayesian recalculation at the detector level (Vishwanathan & Varshney, 1997). In other words, instead of having the detectors (or group members) share their information to a fusion center (or what can be thought of as a non-member judge), the group members should share the information with each other, allowing the local recalculation of each members' response criterion (except for the selected speaker, who is uninfluenced by their own declaration). Research confirms that Bayesian recombination is optimal for group performance when performed by the detector itself (individual group member) (Pete et al., 1998; Shi, Sun, & Wesel, 2001; Vishwanathan & Varshney, 1997). This method of Bayesian updating by the individual group members is most powerful when the voting

takes place in parleys (voting rounds similar to deliberation) because group members are given multiple opportunities to adjust their own criterion (Swaszek & Willett, 1995). These findings mesh well with other studies of information aggregation by individuals which show that a Bayesian combination of old and new information is the best way for individuals to update their own beliefs, especially when information about the source of the information is used (Birnbaum & Mellers, 1983; Birnbaum & Stenger, 1979; Birnbaum, Wong, & Wong, 1976; Birnbaum & Zimmerman, 1998; Morris, 1974, 1977; Sorkin, Luan, & Itzkowitz, 2004). We will use this technique to describe how individual group members should update opinion and how this recalculation affects both individual and group performance.

The SDT Model of Group Deliberation

Signal Detection Theory (SDT) was first applied to psychology by Wilson Tanner and John Swets (1954), and is probably best known for its ability to describe various elements of sensory and perceptual behavior (Macmillan & Creelman, 2004; Wickens, 2001). Most recently, the Sorkin lab has used the SDT methodology to model and describe various aspects of individual and group behavior including advisor preference (Luan, Sorkin, Itzkowitz, 2004), information purchasing (Luan, Sorkin, Itzkowitz, submitted), and collective decision making (Sorkin & Dai, 1994; Sorkin, Hays, & West 2001; Sorkin, West, & Robinson 1998; Sorkin et al., 2004). Whereas the previous SDT work on group decision making has focused on single vote decisions (Sorkin et al., 1998, 2001) or on parallel combination of group members' votes (Sorkin et al., 2004), this paper presents a unique extension of the existing theory to the issue of sequential group deliberation and decision making.

Before presenting the model, it is crucial to note some of its critical assumptions. First, it is assumed that the group members are rational decision makers and have an interest in obtaining a group decision. Secondly, it is assumed that the group members present their decisions to the group in a discrete, binary format. Third, group members utilize the information provided through deliberation to update their own predictions about the nature of the stimulus through a Bayesian combination of the available public information and their own private information obtained through their observation.

In order to best explain our model of group interaction and information combination, we will present it through the context of the deliberative task employed in the simulations and by our human subjects. Because we are interested in how the characteristics of the individual members affect final group performance, we first define two factors that drive individual decision performance. First, for all of the group members, discrimination expertise is given by the detection parameter, d', which is the normalized separation between the means of the two (assumed) normal distributions, as shown in Figure 1-1. Second, all group members possess an individual decision criterion, β_i. Swets, Tanner, and Birdsall (1958) show that if the decision maker is optimal, then this criterion is a function of the prior odds and individual's utility structure:

$$\beta_i = \frac{V_{correct-rejection} + V_{false-alarm}}{V_{hit} + V_{miss}} \cdot \frac{p(noise)}{p(signal)} \qquad [1]$$

Here, $V_{correct-rejection}$, $V_{false-alarm}$, V_{hit}, and V_{miss} represent a member's specific utilities for each of the possible decisional outcomes and $\frac{p(noise)}{p(signal)}$ represents the prior odds ratio. Taken together, d' and β fully define an individual group member's performance in the

individual decision task, both describing total decisional accuracy as well as response bias.

The deliberative process begins with each group member (i) receiving a private, noisy observation (x_i) drawn from either the signal or noise distribution. While each group member's exact observation is assumed to be independent of all other group members' observations, each member's specific observation is selected from the same distribution. After the presentation of this observation, it is assumed that each member of the group forms a likelihood ratio, $l(x_i)$, based upon her individual observation. This likelihood ratio is defined by two distributions, $f(x|N)$, the likelihood that the observation was from the noise distribution, and $f(x|S)$, the likelihood that the observation originated from the signal distribution. Each member then decides the origin of her private observation. To make this decision, each group member's calculated likelihood ratio is compared to her individual decision criterion where the decision rule is:

Respond "Signal" if $l(x_i) \geq \beta$, else Respond "Noise" (also see Figure 1-1).

Following this personal observation and decision, group members begin the process of deliberation. Here deliberation is defined by group members proclaiming their individual decisions to the group in a sequential manner. The rule governing this sequence is known as the *response protocol*. This response protocol selects a group member according to a specific mechanism. The selected member, known as the speaker, then publicly declares her private decision to the group. In addition to the speaker's disclosure, the other group members are given information about the speaker's d' and β. This new information allows the rest of the group to modify their own individual decision criteria. While we describe the method for members to adjust their individual decision

criterion, this method is computationally equivalent to members' recalculation of likelihood ratios.

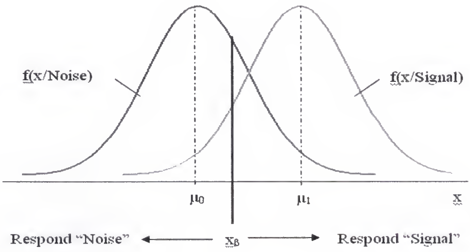

Figure 1-1. Graphical representation of an individual's decision rule.

Each member besides the speaker adjusts in the following manner. First, in order to recalculate one's individual decision criterion any adjustment depends on revision of the prior odds ratio (because the group members' individual utility structures remain constant throughout the task). By using both the speaker's vote and the information provided about the speaker's decisional characteristics, the individual calculates an improved prior odds ratio by performing the following calculation:

$$\beta_{post} = \beta_{initial} \cdot \frac{P(n \mid Vote_{spk})}{P(s \mid Vote_{spk})} \qquad [2]$$

where

$$P(n \mid "Signal"_{spk}) = \frac{P("Signal"_{spk} \mid n)P(n)}{P("Signal"_{spk} \mid n)P(n) + P("Signal"_{spk} \mid s)P(s)} \qquad [3]$$

$$P(\text{"}Signal\text{"}_{spk} \mid s) = \int_{x_{c_{spk}}}^{\infty} f(x - d'_{spk})dx \qquad [4]$$

$$P(\text{"}Signal\text{"}_{spk} \mid n) = \int_{x_{c_{spk}}}^{\infty} f(x)dx \qquad [5]$$

$$x_{c_{spk}} = \frac{\ln \beta_{spk}}{d'_{spk}} + \frac{d'_{spk}}{2} \qquad [6]$$

where *spk* is the current speaker and *f(x)* is the unit normal distribution.

In summary, after each speaker declares her vote to the group, each group member (besides the speaker) recalculates her individual decision criterion in order to utilize the new information received. Using these revised estimates, the group members engage in a secret ballot. These votes are then tabulated and the total is compared to the specific majority required for consensus, known as the group's response criterion. If the number of votes exceeds this criterion then the group vote is recorded. If consensus is not reached, then the process iterates similar to the technique described by Swaszek & Willett (1995), with group members continually speaking to the group, and holding subsequent polls, until either consensus is reached or the group exceeds a specified number of votes. If it is indeed the case that the group exhausts the allotted number of votes, the group vote is reported as "hung." A schematic representation of the decision and deliberation process is given in Figure 1-2.

Response Protocols

As stated earlier, we are quite interested in discovering the nature of the relationship between response protocol and group performance (both in terms of accuracy and deliberation speed). A previous examination of this question under the context of distributed detection has shown that due to the variability inherent in the decision

scenario, there is no one optimal rule, but that specific solutions can be tailored to groups based on the characteristics of the group members (Vishwanathan & Varshney, 1997). So while we forsake the goal of a universal organizational structure that maximizes performance for all groups, we hope that by mapping some of the most plausible and ubiquitous response protocols, we will be able to determine which organizational structure should be utilized in specific situations.

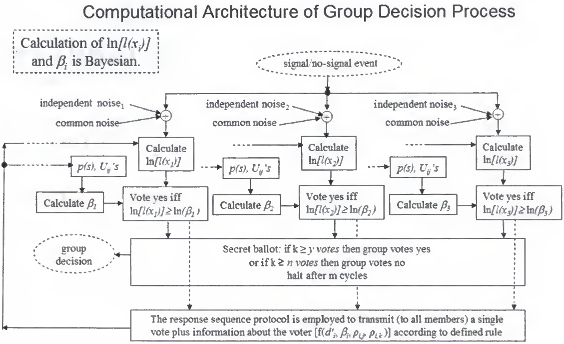

Figure 1-2. Schematic representation (Sorkin, 2005) of the decision and deliberation process.

Because our goal was to maximize accuracy while minimizing deliberation length, we suspected that response protocols that contained the most information should be presented first because it was generally thought that opinions that are presented early in the deliberation process hold a higher value to the group members than opinions presented later in discussion (Carlston, 1977; Davis et al., 1988; French, 1956). Thus, presenting the most informative opinions to the group early in the discussion maximizes

the value of these opinions, potentially swaying the group to a quick, and hopefully accurate, decision. To this end, we tested two protocols that were meant to maximize the information present from the first few speakers to the group.

Because previous research has shown that the most informative group members are the most expert group members (Einhorn & Hogarth, 1975), we examine the effect of ordering speakers based upon their expertise (as defined by *d'*). This response protocol, which we will call the "Max *d'* Rule" defines the speaking sequence by having the group members present their opinions to the group in descending order of their *d'*. Mathematically stated, member 1 would be selected to speak to the group first if the following statement is true.

$$d_1^{'} > d_i^{'} \quad \forall \quad i \in Group \qquad [7]$$

Aside from member expertise, it is also possible to measure unique information through a different metric. Indeed, the "amount" of information an individual possesses is clearly dependent on the initial value of the individual's decision criterion coupled with the strength of their private observation. When speaking order is selected based upon the absolute difference between members' observations and decision criteria, they utilize the response protocol known as the "Difference Rule." Mathematically stated, the Difference Rule would select group member 1 as the speaker if the following were satisfied:

$$\left|l(X_1) - \beta_1\right| > \left|l(X_i) - \beta_i\right| \quad \forall \quad i \in Group \qquad [8]$$

From another viewpoint, it may be useful to think of this member as the surest of her decision. Individuals that receive observations that are "close to" their decision criterion are likely to feel some degree of uncertainty in their decision whereas group

members that receive observations that are drastically different from their individual decision criterion are likely to feel very confident that their opinion is correct. Indeed, studies have shown that confident group members tend to participate in discussion more and have a larger influence on the group's final decision (Sniezek & Henry 1989). Carlston (1977) also found that more confident members of the group were more likely to voluntarily participate first in the deliberation process. For both of these response protocols, placing the most informative opinions to the group early in the deliberative process should hasten the group's decision due to the power of Bayesian updating.

In stark contrast to the aforementioned response protocols, which attempt to maximize accuracy and minimize deliberation length through the presentation of highly diagnostic cues early in the deliberative process, we are also interested in response protocols that may prolong deliberation and thus increase the amount of public information. In their study examining the effect of differing group decision criteria, Guarnashelli, McKelvey, and Palfrey (2000) noted that there was a direct positive correlation between the amount of deliberation before the group decision and the accuracy of the group decision. Guarnashelli et al. (2000) prolonged deliberation length by changing the group's decision criterion from a simple majority rule to a unanimous decision rule; though interventions that change the group's response protocol should have similar effects on group accuracy. In order to test this hypothesis, we chose a response protocol that would present opinions from the least informative group members, those with the lowest d', first in deliberation. Thus, this response protocol, which we will call the "Min d' Rule," sequences the speakers in ascending order of their individual d'.

Mathematically stated, group member 1 will speak to the group first if the following condition is satisfied:

$$d_1^{'} < d_i^{'} \quad \forall \quad i \in Group \qquad [9]$$

Finally, to see if changes in response protocol were significant at all, we include a nominal response protocol that randomly orders the group's speakers. For simplicity, this response protocol will be known as the "Random Rule." For purposes of our analysis, this response protocol will represent the null hypothesis that there is no effect on group performance as a function of the response protocol employed. Indeed, it is quite possible that in groups in which little difference exists between the members' abilities and/or individual decision criteria, there would be no advantage to presenting more informative decisions earlier or later in the decision process due to the marginal differences in speakers' information.

Group Decision Criteria

While we are mainly concerned with how different response protocols affect group performance, it is critical that we note the differences in group performance due to changes in the group's decision criteria. For groups using a single-vote, non-deliberative process, multiple studies have confirmed that a simple majority rule group decision criteria yields the best group performance (Hastie & Kameda, 2005; Sorkin et al, 1998; Sorkin et al, 2001). However "robust" these majority rules are, the simple majority rule does not maximize group performance in groups engaging in a deliberative process. Previous inquiries report that deliberative groups using unanimous decision criteria have both higher group accuracy and longer deliberation lengths (Guarnashelli et al, 2000; Itzkowitz, 2003; Sorkin et al, 2004). Because of the uniformity of this result across

experiments (and research labs), in the studies presented in this paper, we test the majority rule decision criterion only once, using nominal groups (thus serving as a benchmark with which to measure the performance of the other response protocols). We choose to focus almost exclusively on groups using a unanimous group decision criterion because it increases the length of time needed for decision. As seen in Itzkowitz (2003), groups that used the majority rule decision criterion rarely required more than one speaker before consensus was reached, resulting in minimal observable differences between the various response protocols tested. Thus, by implementing the unanimous group decision criterion, we are able to increase deliberation length which will hopefully allow us to determine differences in performance as a function of the response protocol used.

CHAPTER 2
EXPERIMENT 1

Before we investigated how we could maximize group performance, we needed a descriptive account of how groups behaved without any (or minimal) intervention. Gaining a better understanding of natural group behavior would let us determine if there was any value to implementing various response protocols and could also inform us which of the aforementioned protocols could hold the most potential to improve group performance. In this experiment, we attempted to determine the answers to three crucial questions: what response protocol (if any) best matches the sequences used in natural deliberative groups? What is the accuracy of these groups? And, how quickly do these groups reach decision?

Although we desired to have a group free of all interventions, the nature of deliberative groups undermined this goal. Because groups engaged in sequential deliberation must follow some preordained order, some sequencing mechanism had to be imposed onto the group. Through some deliberation of our own, we decided that the group's speaking order should be ordered by its members' initial responses to the stimulus. That is to say, the speaking order was determined by sequencing each member's observational response times in ascending order. Mathematically, group member 1 would be the first speaker if the following is true:

$$rt_1 < rt_i \quad \forall \quad i \in Group \qquad [10]$$

This response protocol is known as the "Natural Rule." While the decision of linking member's willingness to participate based on the speed of her response may seem

obvious, we recognize that speed of response to the observation is in many ways related to, and is a reflection of, an individual's confidence in her initial decision. However, having the most confident individuals declare their opinion to the group early in deliberation was consistent with previous findings that members with high individual confidence in their opinions (in an information based task) were the most likely to be among the first speakers in group deliberation (Carlston, 1977; Sniezek & Henry, 1989; Zarnoth & Sniezek, 1997).

Even though this sequencing technique is consistent with obtained speaking orders from groups with voluntary participation, we recognize that the Natural Rule makes it likely that speaking order will be a function of the group members' individual characteristics and/or observations. We isolate three influences on an individual's reaction time. First, a group member that receives a particularly distinct observation (either Signal or Noise) may respond more quickly due to the nature of her observation alone. Second, response time could be related to the difference between a member's observation and her individual decision criterion (similar to the difference rule). As we previously stated, these members hold unique information that is distinct to them, but not necessarily the result of an overly strong or weak observation. Third, because member ability is dependent on the variance of the display, members that are less able will receive "noisier" observations than those with higher ability. Thus, expert members could be more likely to respond more quickly to the initial stimulus than members with lower ability. We predict that the response times will most likely correlate with at least one of these phenomena, with the difference between the observation and individual decision criterion being the most directly correlated to response time.

Method

Individuals performed a visual signal detection task in groups of 6 members. On each trial, the group members were presented with a multi-element visual display consisting of nine analog gauges (Figure 2-1). This experimental technique has been used with experimental groups in the past (Sorkin et al., 1998; Sorkin et al., 2001).

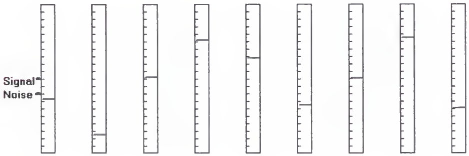

Figure 2-1. An example of a typical observation. The horizontal markings indicate the value of each gauge. These values represent individual, independent draws from the chosen distribution.

Following this initial observation, the group began the process of deliberation. During deliberation, member's opinions were presented to the group using the displays first used in Luan et al. (submitted) (Figure 2-2). At the cessation of deliberation (either due to consensus or due to time limits), all of the group members were provided with feedback about the correctness of the group decision, their individual payoff on that trial, and their cumulative payoff to that point.

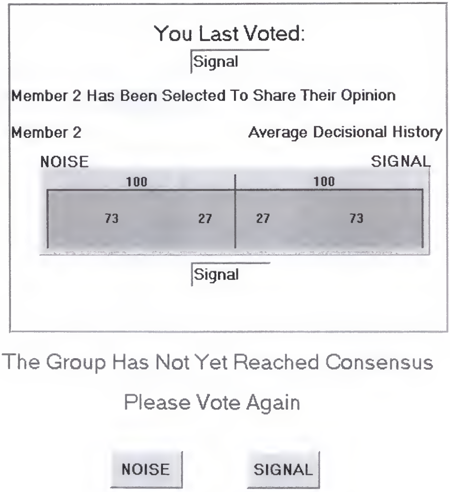

Figure 2-2. Actual display shown to group members that were not the chosen speaker.

Participants

Fifteen University of Florida students participated in this experiment. All subjects had normal or corrected to normal visual acuity. Participants were paid an hourly wage of $6.15 per hour (in accordance with state law) and received a bonus that depended on both the subject's and the group's accuracy, with the group's decision being twice as valuable as the individual decision. During experimental sessions, this bonus, which was equivalent to one cent per point, ranged from $6.00-$11.00 per hour, with most subjects earning $8-$9 per hour (the bonus was halved to .5 cents per point during the training

sessions). This payment structure attempted to ensure that individuals were motivated to make quality decisions both at the individual and group level.

During training, one participant left the study due to scheduling conflicts. This left 14 viable participants for the group study.

Apparatus and Stimuli

Stimulus presentation and response collection were controlled by six 486 computers (clients) that were arranged into a local area network with a Pentium 3 computer (master). The master computer was responsible for synchronizing the clients, generating the stimulus events, employing the response protocols, and testing for consensus. All experiments took place in a quiet, well lighted laboratory. Participants made responses using the computer mouse.

Each trial began with the presentation of the observation. Figure 2-1 gives an example of a generated stimulus event. Each gauge had a series of tick marks on the left side, dividing the gauge into intervals. Each one of these intervals represented one half of a unit of height resulting in a total height of 10 for the entire gauge. During the observation, a red horizontal line simultaneously appeared upon each gauge, indicating the value of that element. The value of each display element was independently generated from one of two equal variance Gaussian distributions, the signal distribution or the noise distribution. So while each element could have a different value, all values were samples from the same distribution. The mean of the signal distribution was higher than the mean of the noise distribution ($\mu_{Signal} = 5, \mu_{noise} = 4$). For all trials, there was an equal probability for either distribution to be chosen. While each member received an

independent observation, all participants' observations were drawn from the same distribution (either signal or noise).

An observer's ability in the decision task was a function of the variances of the signal and noise distributions. Based on previous experiments with this visual detection task (Montgomery & Sorkin, 1996), we knew that subjects only performed at about 75-80% of the predicted *d'* (given by the display's signal-to-noise ratio). For our expert group members, we desired *d'*=1.5, so for these members, the variance of both distributions was set at $\sigma = 1.5$. Likewise, we set the variance of the distributions at $\sigma = 3$ for the group members who we wished to obtain *d'*=0.75.

Stimulus presentation lasted for 300ms in all conditions. The presentation of the stimulus was preceded by a 200ms blue fixation cue that appeared in the center of the nine graphical elements. Each presentation was followed by the appearance of a white masking screen, which appeared for 200ms. Following this screen, the subject was given the opportunity to respond by using the mouse to click on either the "Signal" or "Noise" buttons, according to their decision.

Participants responded to these observations individually and these responses were sent to the master computer, which then determined the speaking order based on the group members' response times to the initial stimulus. The master computer then sent the first speaker's vote and decisional characteristics to the client computers, where it was displayed to all of the group members except for the speaker. Recall that the display of speaker vote and decision information is shown in Figure 2-2.

In this display, the relative size of the appropriate rectangular areas corresponded to the magnitude of the given speaker's predicted correct rejection, false alarm, miss and

hit rates (from left to right); the total size of the correct rejection and hit areas (the outside areas) indicated the speaker's predicted percent correct rate. By presenting these metrics (and the speaker's vote) to the individual group members, each was provided with all of the information needed to recalculate their opinions based on Equations 2-6. Although these measures were presented to the group members as the "Speaker's Average Decisional History" for their last 200 trials, they were actually predictions of the member's actual performance in this decision task based upon the variance of the speaker's Signal and Noise distributions as well as by the speaker's individual utility structure (as shown in Equation 1).

Along with the presentation of this speaker's vote, the other group members were again given the opportunity to vote either "Signal" or "Noise." The speaker was not given the opportunity to change her vote. Following this vote, the master program again tabulated the votes and compared this tally to the group's consensus requirement. If this criterion was met, then deliberation ceased and feedback was given to the participants relating the correctness of the group's decision, individual payoff on that trial, and cumulative payoff to that point. Otherwise, the master computer selected the next speaker and the process of deliberation continued.

Experimental Procedure

Participants received extensive individual training before engaging in the group decision task. Training for all of the participants began with 900 trials performing an individual decision task. This enabled them to become familiar with the task in general and with the various utility structures that could be introduced during the experimental phase. The purpose of this training was to ensure that the participant could behave in a

manner consistent with the condition to which they would be randomly assigned during the experimental phase.

Following this training, subjects received 900 trials performing a more extensive individual decision task. As before, the participants made binary decisions following their own observations, but now, they were presented with the opinion from a simulated "group member." The process by which these group members (and their decisions) were created using the same method used to create the simulated advisors as Luan et al. (submitted) and the individual group members in Itzkowitz (2003). The vote from these simulated members was displayed along with a representation of their decisional characteristics using the methods previously described. This additional training ensured that the participants understood how to interpret and use the information provided in the displays. Specifically, the simulated members exactly represented all possible combinations of group members that the individual would encounter in the group setting.

Following the individual training, the subjects were placed into two groups of seven members each. While we only required groups of six members, we assigned a seventh member to each condition to serve as an alternate if a participant could not attend an experimental session. During the experimental sessions, if all members were present, a member was randomly excused from the session. This same random draw was used to select the member characteristics for each member. Because the resulting bonuses would differ depending on assignment of expertise and bias we wanted to ensure that each member had equal probability of assignment to each condition.

Both of our experimental groups had the same composition. Each group had three members with high expertise d'=1.5 and three members with low expertise d'=0.75. Of

the three members at each expertise level, there was a liberal member ($\beta = .6061$, c= -.5), a neutral member ($\beta = 1$,c= 0), and a conservative member ($\beta = 1.65$,c= .5). Liberal members were given a utility structure that made it more likely that they would respond "Signal," and conservative members' utility structures made it more likely for them to respond "Noise." These utility structures were only present on the first vote. Whereas the neutral group members' had equal payoffs for all four of the possible decisions, group members assigned to either the liberal or conservative conditions received higher penalties for specific types of incorrect answers. For the liberal group member, this was done by penalizing the subject more for misses, or instances of signal occurrence that they deemed "Noise". The same strategy was used for the conservative members, except that they were penalized for false alarms, or instances where they declared "Signal" and the noise distribution had been chosen. While these specific penalties would actually lead to the more extreme betas $\beta = .571$ (c= -.56) and $\beta = 1.75$ (c= .56) previous experience indicated that individuals may need more extreme penalties to drive them to desired bias levels due to a tendency to anchor, so we expected actual performance to be near the desired bias points. Table 2-1 gives the expected characteristics of the group's members based on the variance of their distributions and their individual penalties. The utility structure for the group's decision was symmetric. Each decision had a payoff or penalty of 10 points, depending on the correctness.

Table 2-1. Desired member characteristics for the experimental trials. Note that we have given c to represent the individual decision criterion for group members.

	Member 1	Member 2	Member 3	Member 4	Member 5	Member 6
d'	1.5	1.5	1.5	0.75	0.75	0.75
c	-0.5	0	0.5	-0.5	0	0.5

Each group performed 200 trials in the experimental session, which lasted between 2 ½-3 hours. Because we were wary of our participants succumbing to fatigue, a 5-10 minute break was taken after every 50 trials. While the subjects were told that their response times were being monitored, they were told not to sacrifice accuracy for speed. Participants were informed of the group's decision criterion, but not of the response protocol. These groups performed under the requirement for unanimous consensus.

Results and Discussion

Manipulation Check

It was assumed that each member was accurately represented by the display shown to the group when that member was the speaker. Because these displays remained indifferent to subjects actual performance during the task (except for the display of their actual vote), it was crucial that the group members behaved similarly to their desired decisional characteristics. During the deliberation process, group members' expertise and individual decision criterion changed as a result of hearing new information from each speaker. Thus, the only way to measure group members' observational performance was to examine their response to the initial stimulus. Tables 2-2 and 2-3 show the actual, "first vote" individual member performance for both experimental groups. In both groups, members achieved decision accuracies and response biases very close to the specified amounts. Though the actual performance varied slightly from the desired levels, there is little reason to believe that this had a substantial impact on the groups' overall performance in the deliberative task. Previous work has shown that group members were able to ascertain other group members' actual abilities, even when these abilities were substantially different from their initial expectation (Baumann & Bonner, 2004). So, even if the displays were wildly incorrect indicators of the speaker's actual

decision performance, due to the feedback provided and the large number of trials, members would eventually learn each other's true characteristics. Moreover, Henry (1993) found that knowledge of member ability could still be determined sans feedback. Because the differences between the displayed information and participants' actual behavior were not substantial in our experiment, we believe that these differences had no significant effect on the other group members' ability to correctly utilize the information provided to them in the displays.

Table 2-2. Initial vote member accuracy (*d'*) for both groups using the Natural response protocol and the unanimous consensus.

	Member 1	Member 2	Member 3	Member 4	Member 5	Member 6
Group 1	1.61	1.35	1.54	0.72	0.82	0.67
Group 2	1.42	1.53	1.29	0.78	0.59	0.88
Mean	1.51	1.44	1.41	0.75	0.71	0.78
Std Dev	0.13	0.13	0.18	0.04	0.16	0.15

Table 2-3. Representation of group members' response bias (c) on their first vote. Members with a negative value of c are the "liberal" members and members with positive c are the "conservative" members.

	Member 1	Member 2	Member 3	Member 4	Member 5	Member 6
Group 1	-0.59	-0.13	0.46	-0.50	-0.10	0.56
Group 2	-0.48	0.02	0.54	-0.39	0.16	0.41
Mean	-0.53	-0.06	0.50	-0.44	0.03	0.48
St Dev	0.08	0.11	0.05	0.08	0.19	0.11

The Natural Response Protocol

Recall that the Natural response protocol sequenced the groups' speakers in ascending order of their response times to their initial observations. We attempted to determine which, if any, of the possible response protocols had the most significant influence on these sequences. Because the two experimental groups were not statistically different from each other in terms of individual member or group performance, we combined the groups to perform these analyses. Figures 2-3, 2-4, 2-5, and 2-6 show the

results of these analyses. In all cases, we computed a least squares regression to obtain r^2.

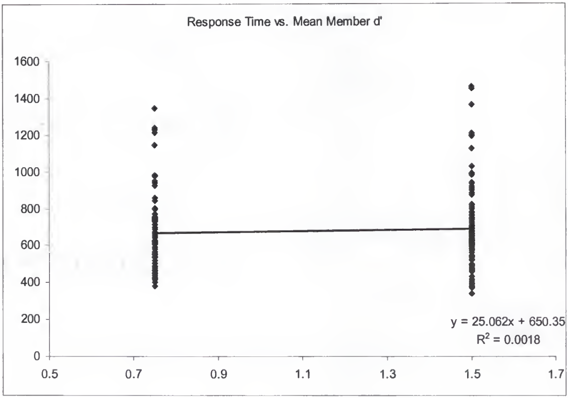

Figure 2-3. Response time of the speaker as a function of the member's *d'*. Whereas the individual data points represent the response times recorded, the black line is a presentation of the least squares function. The amount of variation in response time caused by variation of the speaker's *d'* is given by r^2.

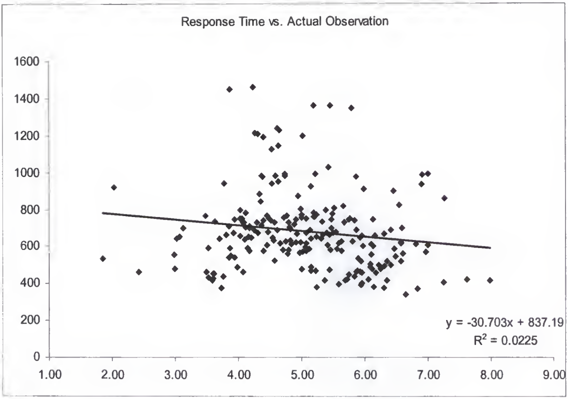

Figure 2-4. Response time of the speaker as a function of the speaker's mean observation. Whereas the individual data points represent the response times recorded, the black line is a presentation of the least squares function. The amount of variation in response time caused by variation of the speaker's mean observation is given by r^2.

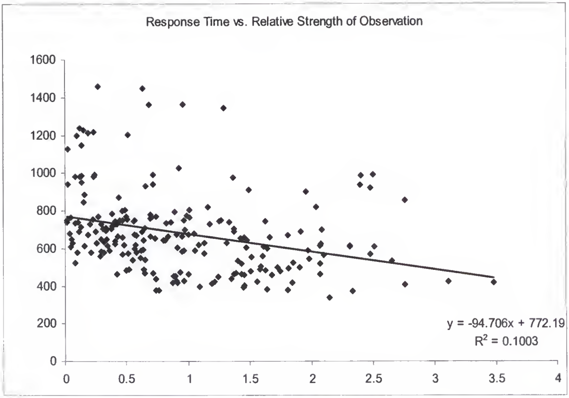

Figure 2-5. Response time of the speaker as a function of the relative strength of each speaker's observation. Larger numbers of strength represent observations that were far from the display's midpoint. Whereas the individual data points represent the response times recorded, the black line is a presentation of the least squares function. The amount of variation in response time caused by variation of the relative strength of the speaker's observation is given by r^2.

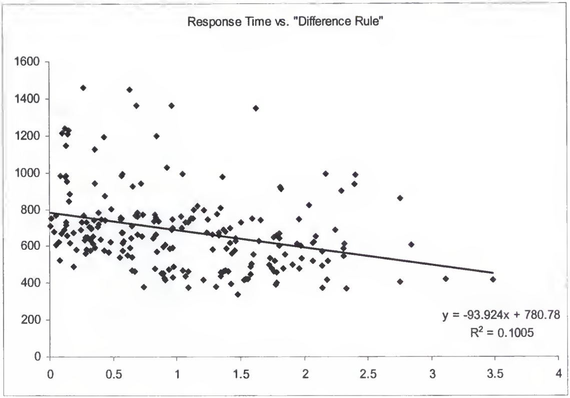

Figure 2-6. This figure shows the response time of the speaker as a function of the absolute value of the difference between a speaker's mean observation and their individual response criterion. Whereas the individual data points represent the response times recorded, the black line is a presentation of the least squares function. The amount of variation in response time caused by variation of the difference between the speaker's individual decision criterion and their observation is given by r^2.

Each figure contains the linear regression function and r^2. Of the possibilities tested, both the relative strength of a speaker's observation and the difference between the speaker's criterion and her observation had a larger influence on response time than either the ability of the member or the exact mean of the observation received. It is important to note that none of the measures tested contributed a large amount to the variation in the speakers' response times. There are two possible explanations for the observed results. First, response time to the initial stimulus may not be related to confidence at all. Instead, the observed differences could have been due to individual

members' variation in response time. Thus, equating response time to member confidence may have been inappropriate. Second, it is possible that response time was related to the speaker's confidence, but that observational strength (in whatever form) and member ability were not large sources of this confidence. If the group was to participate in an exchange of information relating continuous estimates of likelihood (instead of the binary ones used in this experiment), we would predict there to be a stronger relationship between the members' initial rating and their confidence since the rating itself is a type of confidence indicator. This is certainly an area of investigation that warrants future study.

Group Performance

Group accuracy

We calculated the accuracy of both groups based on the correctness of the binary consensus decision made on each trial. The result of this analysis for both groups is given in Figure 2-7. Following the procedure given in Macmillan & Creelman (2005) we computed the variance of *d'* for each group using the following equation:

$$Var(d') = \frac{HR(1-HR)}{N_{Signal}(\phi(HR))^2} + \frac{FA(1-FA)}{N_{Noise}(\phi(FA))^2} \qquad [11]$$

Where HR is the hit rate, FA is the false alarm rate, N_{Signal} is the number of signal trials, N_{Noise} is the number of noise trials, and ϕ is the normal density function. These measures of variance were used to determine the standard error of *d'*. While these measures are smaller than the standard error of *d'* they better represent the variation of pooled *d'*. For group 1, d'=2.04 (SE=.22), and for group 2, d'=2.01 (SE=.21). Average group performance was d'=2.02 (SE=.1). As shown by the small standard error of *d'*,

performance was relatively stable between the two groups. We also calculated an ideal group based on Green and Swets (1966) technique (because the members were

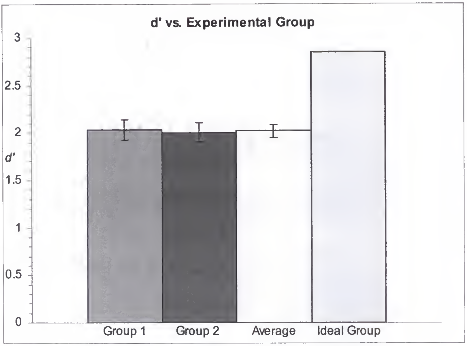

Figure 2-7. Group *d'* by experimental group, Aggregate group *d'*, and *d'* for the ideal group (created from average achieved member first vote performance).

independent, we did not use the Sorkin & Dai method). Because the ideal group assumes perfect information flow and use, it is not surprising that both experimental groups had lower decision accuracy than the ideal group. The ideal group shown above was calculated using the mean of the observed member characteristics of both groups yielding d'_{ideal}=2.85. Using the ideal group as a benchmark, we calculated the efficiency for the combined groups (η=.50). As expected, both groups had higher accuracy than any of their constituent members, but accuracy was lower than the ideal group. This is certainly understandable given the nature of the deliberative task. Because deliberation ceased before all group members were given the opportunity to present their opinion to the group, there was no way the group could have used all of the information available. As

shown in the discussion below, groups rarely heard more than two public declarations before consensus was met. Similarly, the group performed better than its best member because each speaker increased the accuracy of the other group members. Whereas all group members began deliberation with the accuracy shown in Table 2-2, the post deliberation accuracy for these members was the same as d'_{group}, a significant increase for all of the members.

Bias

We were also interested in how deliberation influenced the groups' decisional bias. After recording the achieved *c* for each group, we again used Macmillan and Creelman (2005) to calculate the variance of *c* for each group using the equation

$$Var(c) = .25(Var(d')) \qquad [12]$$

where var(*d'*) is given by equation 11. Using these variances, the standard errors of *c* were calculated. The post deliberation group bias and associated standard errors are given in Table 2-4. As is clear from the standard errors shown, there was no significant difference in *c* between the two groups. Because both groups began deliberation with a mean *c* of zero, and because members with biases were only rewarded (or penalized) for them on their initial vote, it follows that bias would trend towards a neutral state.

Table 2-4. Post deliberation response *c* bias and associated standard errors for experimental groups using the Natural response protocol and a requirement for unanimous consensus. The averaged decision bias of both groups is also given.

	Group 1	Group 2	Average
c	0.050	-0.030	0.010
S.E.	0.107695	0.10388	0.0882

An analysis of members' post-deliberation bias revealed that members' final decisions were about neutral. Thus, for the group members that began deliberation with

either a (substantial) liberal or conservative bias, this bias was all but mitigated by their final decision. This replicated the findings of Itzkowitz (2003), which determined that imposing a unanimous group decision criterion served to "debias" groups, even when the groups did not begin with a neutral response bias. We believe that there were two main sources of this effect. The combination of a symmetric utility structure for group consensus decisions and the added incentive for unanimity (groups were not paid for hung trials), both contributed to this effect.

Deliberation length

Aside from measures of accuracy and bias, we also calculated the length of the groups' deliberations. Here, the average number of group polls taken represents the "time" it took the group to reach consensus. Because the group participated in a secret ballot after each speaker, this number also represents the number of speakers that had been given the opportunity to make public declarations of their opinions before the groups made their final decisions. The average deliberation lengths and their associated standard errors for both groups are given in Table 2-5. There was no significant difference in deliberation length between the groups. For both groups, a minimum of one speaker was needed before the groups could reach consensus. The maximum deliberation time in either group was seven votes (group 1 had two of these instances). There were no hung trials in either group.

Table 2-5. Average deliberation length and associated standard errors for groups using the Natural response protocol and the unanimous decision criterion.

	Group 1	Group 2	Average
Mean Votes	2.375	2.289	2.332
S.E.	1.050	1.021	1.036

The speed at which the groups reached consensus is consistent with other studies of group deliberation. Previous simulations of group deliberation demonstrated the

power of each member updating their individual decision criterion in a Bayesian manner (Itzkowitz, 2003; Sorkin et al., 2004). Unfortunately, previous models of group judgment and discussion all but neglect the issue of deliberation length. While this study allowed us to determine that group members reached consensus rather quickly, because we only studied one group, it was impossible to tell if there was a speed-accuracy tradeoff within deliberative groups.

CHAPTER 3
EXPERIMENT 2

In Experiment 1, we determined the behavior of groups that used the Natural response protocol. This allowed us to understand the performance and processes that governed deliberation when a minimal intervention was applied to the group's organizational structure. However, the inability to simulate group members' response times meant that it was impossible to compare the actual group performance to the performance of deliberative groups employing the normative SDT model.

In this experiment, we tested groups, both simulated and actual, to determine performance under a variety of response protocols and decision criteria. This enabled us to make comparisons between normative and actual performance of deliberative groups. Previous research using simulated groups concluded that differences in performance due to group decision criteria were larger than those performance differences caused by response protocol alone, with the largest difference existing between the majority and unanimous decision criteria (Itzkowitz, 2003). However, because the simulated groups in that experiment did not have large variances of member expertise or bias, it is logical to presume that differences in response protocol will have a larger effect when these variances increase (due to the nature of the recalculation process). We predict that groups using the Difference or Max d' rule will reach consensus in the shortest amount of time, because the earliest speakers have the most information to share with the group. It is more difficult to predict which response protocol will produce the highest accuracy for the group. While it is possible that the Difference and Max d' protocol could lead to

highly accurate groups because group members hear the best information first (and the first speakers have larger influence), it is also possible that because these protocols lead to quicker consensus decisions, opportunities to hear additional independent opinions are reduced, therefore decreasing the total amount of public information available before decision. If this loss of information outweighs the impact of group members hearing better speakers earlier in deliberation, then the Min *d'* rule will probably result in the best performance, because it is the most likely to lead to longer deliberation times. With regard to the groups' decision criteria, we expect groups requiring unanimity to have higher accuracy and longer deliberation lengths than groups requiring a simple majority.

Simulations of Deliberative Groups

To test the effects of the various response protocols (and decision criteria) we used a simulation that replicated the task environment and the decisional methods described by the model and used in our experiments with human decision makers. The simulation grouped the experimental trials and conditions in several ways. Most generally, the simulation grouped trials into blocks of 1000 trials. While group composition was static within each block of trials, it did vary between blocks. For each block, the simulated members' ability and individual decision criterion were created by sampling from respective normal distributions (one for *d'*, one for *c*) with the mean and variance controlled by the experimenter. This technique enabled us to more fully investigate groups within a given set of parameters by allowing us to take a sample of different groups (something that is quite difficult with actual groups). Although the group members' characteristics were determined through random sampling, we imposed some limits to the values in order to ensure that no group members' expertise or individual decision criterion was an extreme outlier of the distributions from which they

were selected: d' ranged from 0.2 to 3 and members' c ranged from -2.5 to 2.5. For values that either fell below or exceeded the bounds, the minimum/maximum allowable value was substituted in its place.

Following the creation of the group members, the simulation began the detection task. First, the computer randomly selected either the noise or the signal plus noise distribution (recall that the prior odds are initially even). From this event, the simulation created an observation for each group member. This observation was a sample from a normal distribution which was then weighted by the observers d' (higher expertise of the member resulted in a better observation) and the distribution chosen (the mean of the signal distribution is higher than the mean of the noise distribution).

After these observations had been received by the group's members, a member was selected to disclose their personal decision to the group according to the specified response protocol. After this speaker declared their opinion to the group, the other group members recalculated their own individual decision criterion. This recalculation followed the procedure outlined in equations 2 through 6. As with human subjects, the simulated group members possessed all relevant information regarding the speakers d' and β.

Subsequent to members' recalculation, a poll was taken. If consensus was met, deliberation halted and the next trial began, assigning each member new observations. If the decision criterion had not been satisfied, then the deliberative process iterated, beginning with the selection of a second speaker. The deliberative process was limited to 12 votes. This enabled the group of six to cycle twice. Either reaching consensus and

recording a group vote or exhausting the maximum allowable votes led to the conclusion of each trial within the block.

At the completion of each trial, the simulation recorded measures of the group's performance. Mainly, the simulation recorded the expertise (*d'*) and bias (*c*) of the group. In addition to these typical SDT measures, we also recorded the average number of votes needed to reach consensus and the number of times that the jury was unable to reach decision altogether (hung trials).

Because we intended on using these simulations as a benchmark from which to measure the difference in performance of our human groups, we used the actual means and variances of *d'* and *c* obtained from our human groups to bootstrap the distributions of members' characteristics for each simulated group. While we recognize that this technique results in less valuable comparisons made between simulated groups, it also means that the comparison between the simulated groups and the actual groups using the same organizational structure would be more meaningful. The mean and variance for *d'* and *c* used for each response protocol are listed in Table 3-1.

Simulated groups performed 30 runs of 1000 trials each for each of the following response protocol/group decision criteria combinations: Random-Majority, Random-Unanimous, Max *d'*-Unanimous, Min *d'*-Unanimous, Difference-Unanimous.

Experiments with Human Decision Makers

Method

Participants

Two groups of 7 participants each were formed from the subjects that had participated in Experiment 1. The hourly wage and bonus structure was also the same as Experiment 1.

Apparatus and Stimuli

Experiment 2 used the same facilities and group decision program as Experiment 1. Because the program used allowed the experimenter to select which group decision criterion and response protocol would be used in each session, no changes were made to the either the master or client computers.

Procedure

Group members received extensive training prior to and during Experiment1. No additional training was conducted before Experiment 2. As before, group members were randomly assigned to one of the two experimental groups. During the experimental sessions, a random draw was taken which excluded one participant, resulting in six member groups. This draw also designated each participant's assignment of member characteristics for the experimental session. Individual and group decisions were also rewarded or penalized using the same utility structures described in Experiment 1.

Groups performed the decision task for 200 trials. Whereas groups under the majority requirement for consensus were able to complete these 200 trials in about 1½ hours, groups with a unanimous decision criterion took just over 3 hours. Like before, breaks were taken after every 50 trials to prevent fatigue. While participants were given information regarding the group's decision criterion, they were not given information about the active response protocol. Members were instructed to make the most accurate decision possible.

Results from Simulated and Actual Groups

Just as in Experiment 1, member performance was consistent with the desired levels. Again, there were no significant differences between the sets of experimental groups. In order to decrease the standard error of our analyses, an aggregate group was

created using the data from both individual groups. The aggregate member d', c, and their associated standard deviations for members' initial decisions are given in Table 3-1. The standard deviation given represents the deviance between the individual members of the combined group.

Table 3-1. Aggregate d', c, and their associated variances for group members' performance on their observational decision, given by response protocol and decision criteria.

	Rand-Maj	Rand-Unan	Min d'	Max d'	Difference
Mean Memberd'	1.19	1.09	1.29	1.12	1.16
S.D. d'	0.44	0.25	0.52	0.49	0.38
Mean Member c	0.09	-0.05	0.10	-0.01	-0.06
S.D c	0.35	0.53	0.62	0.50	0.43

For all groups, both simulated and actual, final decision bias of the group was essentially neutral. This was consistent with the findings of Experiment 1 and with previous data that indicated that groups which began deliberation with an average neutral bias had final decisions that were essentially neutral in bias (Itzkowitz 20003, Sorkin et al., 2004). For the sake of brevity, these results were omitted from this discussion.

Group Accuracy

General

We again measured the accuracy of the group's decisions based on their binary consensus vote. The final decision accuracy for both actual and simulated groups is given in Figure 3-1. For actual groups, we again calculated the standard error of d' based on Macmillan and Creelman (2005). Because of the large amount of trials in the simulated groups, we calculated the traditional statistical measurement of standard error.

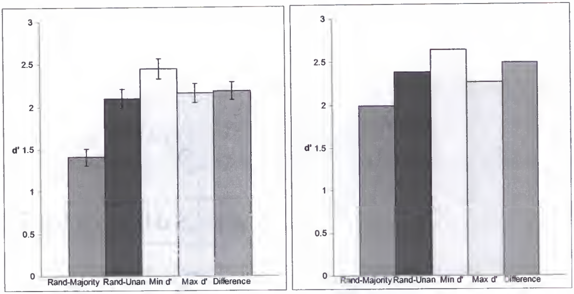

Figure 3-1. Consensus accuracy for simulated and actual groups. The left panel gives both the *d'* and associated standard error for actual groups. The panel on the right shows the achieved *d'* for simulated groups. Standard errors were too small to display for simulated groups (S.E <.02).

For all combinations of response protocol and decision criteria, the actual groups closely matched the simulated groups, although the actual groups were slightly less accurate. For the unanimous groups, the actual groups had efficiencies of η>.77 when compared to respective simulated groups (efficiency for the Random-majority group was η=.50). Note that these efficiencies were higher than we would expect if we were to have made a comparison of the actual groups to the ideal group. Although the simulated group members performed optimal updates of their individual decision criterion, their simulated observations were still noisy. Moreover, the deliberation process itself did not allow all private member information to be disseminated to the group resulting in less than total information pooling. Because the simulated groups used the achieved *d'* and *c* values of the actual groups, the differences in decision accuracy between the two groups were mainly due to differences in member updating. The high efficiencies indicated that our

actual group members updated their own individual decision criteria in a Bayesian fashion, though not always exactly as the formulas described by our model posit.

Response protocols and decision criteria

Differences in response protocol and decision criteria had an influence on consensus performance for all groups, both simulated and actual. For the simulated groups, the difference between each group and every other group was significant ($p<.05$). For actual groups, the only significant difference between groups with different response protocols and decision criteria came from comparisons to the Random-Majority group, which had significantly lower accuracy than all of the other combinations ($p<.05$). For all other human groups, comparisons between them were not significant ($p>.05$). For groups deliberating under a requirement for unanimous consensus, there groups using the Min *d'* rule and unanimous decision criterion had the highest performance in both simulated and actual groups. For both simulated and actual groups, employment of the Max *d'* rule led to the least accurate groups. Except for the Max *d'* ($p>.10$) and Min *d'* ($p<.10$) rules, all of the simulated groups had significantly better performance than the actual groups ($p<.05$).

Deliberation Length

Again, we used average number of votes to decision as a proxy for deliberation length. The mean number of votes and their associated standard errors are given in Figure 3-2. Notice, that in most cases, the actual groups were able to reach consensus in a more timely fashion than the simulated groups. Additionally note for both actual and simulated groups, there was a significant difference in deliberation length caused by the change of decision criteria from majority to unanimous. For actual groups, there were no significant differences between deliberation lengths as a function of response protocol.

However, in general, deliberation times were longest for the Min *d'* rule and shortest for the Max *d'* rule when there was a unanimous requirement for consensus. For the simulated groups, there was a significant difference between all of the groups deliberating with a unanimous requirement for consensus ($p<.05$), except between the Max *d'*-Random response protocol and the Max *d'*-Difference response protocol. In these simulated groups, the Min *d'* response protocol led to the longest deliberations. In descending order of deliberation length, the Min *d'* rule was followed by the Difference rule, the Max *d'* rule, and the Random rule. Differences in deliberation length were not significant between actual and simulated groups using the same organizational structure.

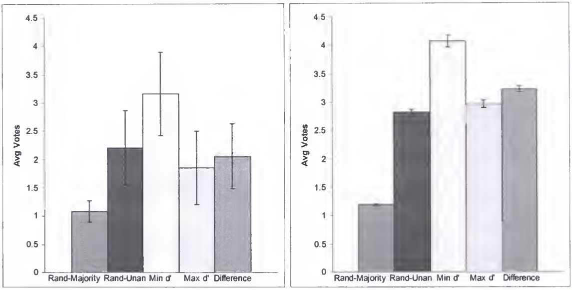

Figure 3-2. Mean deliberation times and associated standard errors for actual groups (left panel) and simulated groups (right panel).

Discussion

Experiment 2 enabled us to answer some essential questions about deliberative groups. Like the findings of Murrell (1977) and Robinson & Sorkin (1985), we found that individuals incorporated additional information from other group members in a Bayesian fashion. This was shown by the high efficiencies that we saw between the

actual and simulated groups. Because all other variables were controlled for, and because the simulated groups used data from the actual groups as a bootstrap, we know that the observed inefficiency of actual groups was due to factors that influenced how members used information from the group's speaker (otherwise performance between the simulated groups and actual groups would have been equivalent) Indeed, although we tried to control for it in every way possible, both normative *and* informational influences exist in deliberation about factual judgments (or states of nature) (Deutch & Gerard, 1955). Thus, although we tried to limit other group pressure to increase the possibility of members behaving in a rational way, a total elimination of these factors may be impossible.

Aside from addressing the question of individual information aggregation and use, we also attempted to determine which response protocols maximized decisional accuracy while simultaneously minimizing deliberation length. No response protocol was able to satisfy both criteria simultaneously. While we were correct in our hypothesis that the Max *d'* rule and Difference rule led to the quickest deliberations in actual groups, these rules did not lead to the best consensus performance. In both simulated and actual groups, the highest decision accuracy was obtained when groups used the Min *d'* rule. In both cases, the Min *d'* rule also led to the lengthiest deliberations.

Clearly, there was a link between the deliberation length and group performance. In order to study this relationship, we examined accuracy as a function of a groups' deliberation time. For both actual and simulated groups, as deliberation length increased, so did consensus performance. Statistical analysis revealed that deliberation length was

indeed a strong predictor of groups accuracy (r^2=.82 for actual groups and r^2=.92 for simulated groups).

Finally, we confirmed that groups using the unanimous decision criterion had better performance than groups using a simple majority rule for consensus. While this had been shown in simulated groups before (Itzkowitz, 2003) and in groups performing a deliberation like task (Guarnashelli et al., 2004), this was the first study to examine this phenomenon using a tightly controlled decision task with strict rules for group deliberation. Taken together with our findings regarding deliberation length, these findings indicate that organizational structures designed to increase the amount of pooled information greatly increased the performance of the group.

CHAPTER 4
GENERAL DISCUSSION

We have proposed a normative model of group deliberation that allows us to determine how organizational structure affects group performance. This model has provided us with a standard against which we have measured actual group behavior. Unlike previous social judgment models of group decision making that rely on probabilistic models of individuals' decisions within the group (Davis, 1973, 1992; Kerr, MacCoun, & Kramer, 1996) our model provides an integrated description of both individual and group decision making under deliberation. Additionally, our model is based on a task where there is a correct answer, so group accuracy can be determined. The lack of a measure of accuracy has been one of the major issues of the social judgment literature (Hastie & Kameda, 2005). The SDT approach allows determination of these judgments based on the characteristics of the individual decision makers. Existing models that provide an integrated treatment of individual and group decision making that include measurable decisions have somewhat different mechanisms for deliberation. These studies either have members share information in a parallel manner (Sorkin et al., 2004; Swaszek & Willett, 1995), or include a decision mechanism outside of the group (Hoballah & Varshney, 1989) and thus do not address the critical issues that govern sequential deliberation. Through our examination of the deliberative task, we have been able to determine how different organizational structures influence the performance of group decisions, both in terms of consensus accuracy and deliberation length.

In experiment 1, we attempted to determine which, if any, response protocols were used in groups where members' participated according to some sort of voluntary structure. Because we desired speaking orders that reflected group members' desire to share their information with the group, the governing response protocol sequenced speakers in ascending order of members' response time to their initial stimulus. Previous research indicated that group members who are certain (or confident) of their decision are the most likely to participate (Carlston, 1977; Sneizek & Henry, 1989). We predicted that this confidence would most likely be related to the observation received. Results indicated that the strength of a member's observation relative to their own individual decision criterion was shown to have an impact on the speed of a member's decision. The deviation of a members' observation away from the center point of the display also had an influence on member response time. Neither of these relationships was particularly strong.

Aside from determining which factors determined speaking orders in voluntary groups, we were also interested in the group performance. First, we found that group deliberation, even with a requirement for unanimity, was relatively quick. This was consistent with the predictions of Swaszek & Willett (1995). Second, group performance was better than that of the best member in the group, providing evidence contrary to previous studies that found that groups rarely perform better than their best member (Gigone & Hastie 1997; Einhorn, Hogarth, & Klempner, 1977). Third, group deliberation with a requirement for unanimity had the effect of "debiasing" the most extreme group members. Whereas two-thirds of the group members began deliberation

with some sort of initial bias (either liberal or conservative), all members consensus decisions were the same, and this group decision had a generally neutral bias.

In comparison to other types of response protocols (such as the ones tested in experiment 2), the Natural response protocol did not fare poorly. Groups using the Natural protocol had the lowest actual detectability of any unanimous group. However, these groups' performance was not significantly different from groups using a Random response protocol. Assuming that response time is a good proxy for the propensity for a member to be chosen to share their opinion with the group, this result indicates that groups participating in voluntary deliberation rather than *discussion* may not warrant additional intervention for slight increases in decision accuracy (at least when there is a unanimous requirement for consensus).

Experiment 2 demonstrated how different organizational structures affected group performance. In this study, normative groups were compared to groups of human members. By making these comparisons we hoped to answer two questions: Do optimal organizational structures exist? And do group members update their individual decision criterion in a Bayesian fashion as assumed by the model? Both simulated and actual groups were tested.

All of the groups tested had higher decisional performance than that of the group's best member. This indicated that groups were using the information provided in deliberation in some way. However, none of the organizational structures tested maximized accuracy concurrent with minimizing deliberation length. In fact, our findings indicated that these two performance metrics have a distinct and powerful

relationship. For this reason, groups using a Min *d'* rule were slightly more accurate than all of the other groups.

Tests of actual groups' efficiency (η) indicated that actual groups performed the decision task nearly as well as the simulated groups. Because we were able to control for other factors, this efficiency represented how Bayesian the individual group members were in their recalculation of their individual decision criterion. We also discovered that there was a difference in efficiency due to changes in the group's decision criterion. Whereas groups requiring unanimous consent had very high efficiencies, groups that required a simple majority had members who did not update in a Bayesian fashion, shown by the lower efficiency than the simulated groups.

To help explain what facets of the majority rule decision criterion would cause its members to act in a manner differently than their unanimous counterparts, we turn to recent studies in economics. It has become popular in the experimental economics literature to study group informational influence under the guise of "information cascades" (Anderson & Holt, 1997; Hung & Pratt, 2001). As defined in these studies, "information cascades" represent situations where individual group members ignore their own private information as the result of their Bayesian integration of previously announced, public information. Hung & Pratt (2001) investigated the role of the majority rule on decision making groups, where individual payoff was dependent on the group's majority decision (much like this study). Groups were also tested under a conformity rewarding group, which was different from our unanimous group in that it rewarded conforming answers more than it rewarded the group's actual decision. It was found that

individuals behaved in a much more Bayesian manner when there was pressure for conformity than when the groups used a simple majority rule.

While the conformity rewarding process was different from the unanimous rule examined in this study, pressure to reach consensus certainly existed in our study. First, individuals received greater payoffs for group decision than for individual decision. Secondly, groups received no payoff whatsoever if group consensus was not met. Third it was obvious from subjects' comments that the unanimous decision criterion occasionally led to uncomfortably long sessions (even with the mandatory breaks), thus there may have been some normative pressure to not be the member that consistently prolonged deliberation by not conforming (although the hourly wage participants received may have tempered this factor somewhat). Although different in payoff structure, it is clear that the unanimous decision criterion rewarded group conformity.

Groups' Bayesian behavior also explains why the Difference and Max d' rules did not lead to higher performance. Group members correctly interpreted the information present, but due to its power, changed their opinions too quickly, resulting in less pooled information and less opportunities for the error checking that occurs as part of the deliberative group process. Thus, because members were Bayesian, presenting expert opinions early in the deliberation led to shorter deliberation length than when less accurate group members were among the first speakers.

Group deliberation is a highly dynamic enterprise. Group performance is a function of group members' characteristics and the organizational structure used to reach consensus decisions. Whereas the former mostly affects individuals' private judgments, the latter affects how these judgments are integrated with the other group members'

information. Our model provides an integrated model for addressing individual and group decision making behavior in deliberative groups and provides a compelling account of actual group behavior in the laboratory. In general, individuals within the groups tested updated their own opinions according to the normative, Bayesian, predictions. Moreover, groups closely matched the predictions of group accuracy and deliberation length provided by the simulated groups. Although our experimental procedure was simplified and did not reflect the complexity of actual deliberative groups, they indicate that groups act in a largely rational way. Moreover, experiment 1 demonstrates that even groups with little intervention have performance close to normative predictions, demonstrating the adaptability of groups where deliberation is governed by voluntary participation, much as in real life (although our participants were extensively trained and well paid).

While these studies have answered some of the basic questions regarding deliberation processes, they raise others worthy of future investigation. First, what is the appropriate way to determine which members speak to the group in naturalistic settings? While there has been extensive work on the topic of speaker participation in discussion (Larson, Sargis, & Bauman, 2004; Stasser, 2000; Stasser & Titus 1987; Stasser & Vaughn, 1996), the decision environment in these studies were substantially different from the deliberative process described here. The determination of natural speaking orders would allow natural deliberations to be compared to a normative standard of deliberative groups, rather than the ideal model. Second, what are the effects of deliberation length separate from that of increased information? By designing a decision environment where a fixed number of speakers presented their opinions to the group

before allowing unlimited polling without speakers, it would be possible to determine if deliberation length is related to the present amount of pooled information. Third, how should groups deliberate when group members' information is highly correlated? In this situation, group members must include information about the speaker's vote, ability, and decisional bias along with information about the degree of dependence between that speaker's information and their own. It is obvious that correlated information should not be weighted as heavily as independent information during the recalculation process, but to what extent do individuals reflect this normative discounting of correlated information? Finally, in the experiments shown, there was no cost to obtaining additional information from the group. This is not the case in many actual decision scenarios, where additional information gathering involves substantial cost, either of time or money. Under what situations should groups choose to purchase more information, and how do these decisions vary based on the quality and cost of the additional information?

The basic issues addressed by our model and the additional questions raised here have real world implications. Because so many important decisions are made by deliberative groups (e.g., juries, military teams, and corporate boards), it is essential that we discover ways to the maximize performance of these groups. Additional applications exist outside of the typical group setting. The techniques described here could also be used to explain market behavior, which is essentially a distributed, deliberative group. In conclusion, our model allows us to show how groups should decide, provides metrics for actual group performance, and may offer prescriptive solutions for inefficient group behavior.

LIST OF REFERENCES

Anderson, L. R. & Holt, C. A. (1997). Information cascades in the laboratory. The American Economic Review, 87(5), 847-862.

Birnbaum, M. H. & Mellers, B. A. (1983). Bayesian inference: Combining base rates with opinions of sources who vary in credibility. Journal of Personality & Social Psychology, 45, 792-804.

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. Journal of Personality and Social Psychology, 37, 48-74.

Birnbaum, M. H., Wong, R., & Wong, L. H. (1976). Combining information from sources that vary in credibility. Memory and Cognition, 4, 330-336.

Birnbaum, M. H. & Zimmermann, J. M. (1988). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. Organizational Behavior & Human Decision Processes, 74, 145-187.

Carlston, D. E. (1977). Effects of polling order on social influence in decision-making groups. Sociometry, 40(2), 115-123.

Clemen, R. T. (1989). Combining Forecasts: A review and annotated bibliography. International Journal of Forecasting, 5, 559-583.

Clemen, R. T., & Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. Risk Analysis, 19(2), 187-203.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. Psychological Review 80, 97-125.

Davis, J. H. (1992). Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples, 1950-1990. Organizational Behavior and Human Decision Processes, 52, 3-38.

Davis, J.H., Stasson, M., Ono, K., & Zimmerman, S. (1988). Effects of straw polls on group decision making: Sequential voting pattern, timing, and local majorities. Journal of Personality and Social Psychology 53(6), 918-926.

Davis, J. H., Stasson, M. F., Parks, C. D., Hulbert, L., Kameda, T., Zimmerman, S. K., & Ono, K. (1993). Quantitative decisions by groups and individuals: Voting procedures and monetary awards by mock civil juries. Journal of Experimental Social Psychology, 29, 326-346.

Davis, J.H., Tindale, R. S., Nagao, D.H., Hinsz, V.B., & Robertson, B. (1984). Order effects in multiple decisions by groups: Demonstration with mock juries and trial procedures. Journal of Personality and Social Psychology, 47(5), 1003-1012.

Deutsch, M. & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. International Journal of Psychology, 51, 629-636.

Einhorn, H. & Hogarth, R. (1975). Unit weighting schemes for decision making. Organizational Behavior and Human Decision Processes, 13, 171-192.

Einhorn, H. J., Hogarth, R. M., & Klempner, E. (1977). Quality of group judgment. Psychological Bulletin, 84, 1, 158-172.

French, J. R. (1956). A formal theory of social power. Psychological Review, 63, 181-194.

Green, D. M. & Swets, J.A. (1966). Signal detection theory and Psychophysics. New York:Wiley.

Gigone, D. & Hastie, R. (1996). "The impact of information on group judgment: A model and computer simulation," In E. H. Witte and J. H. Davis (Eds.), Understanding Group Behavior. Hillsdale, N.J: Lawrence Erlbaum Associates.

Gigone, D., & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. Psychological Bulletin, 121, 149-167.

Grofman, B., & Feld, S. (1986). Determining optimal weights for expert judgment. In B. Grofman & G. Owen (Eds.) Decision Research (Vol. 2, pp 167-172), Greenwich, CT: JAI Press.

Grofman, B., Owen, G., and Feld, S. (1983). Thirteen theorems in search of the truth, Theory and Decision, 15, 261-278.

Guarnashelli, S., McKelvey, R. D., & Palfrey, T. R. (2000). An experimental study of group decision rules. American Political Science Review, 94, 407-423.

Hastie, R. & Kameda, T. (2005). The robust beauty of majority rules in group decisions. Psychological Review, 112, 494-508.

Henry, R. A. (1993). Group judgment accuracy: Reliability and validity of postdiscussion confidence judgments. Organizational Behavior and Human Decision Making, 56, 11-27.

Hoballah, I. Y., & Varshney, P. K. (1989). Distributed Bayesian signal detection. IEEE Transactions on Information Theory, 35, 995-1000.

Hinsz, V.B., Tindale, R.S., & Vollrath, D.A. (1997). The emerging conceptualization of groups as information processors. Psychological Bulletin, 121, 43-64.

Hung, A. A. & Plott, C. R. (2001). Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions. The American Economic Review, 91, 1508-1520.

Innami, I. (1994). The quality of group decisions, group verbal behavior, and intervention. Organizational Behavior and Human Decision Processes, 60, 409-430.

Kaplan, M.F. & Martin, A. M. (1999). Effects of differential status on group members on process and outcome of deliberation. Group Processes and Intergroup Relations 2(4), 347-364.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. Psychological Review, 103, 687-719.

Luan, S., Sorkin, R. D. & Itzkowitz, J. (2004). Weighing information from outside sources: A biased process. Journal of Behavioral Decision Making, 17, 95-116.

Luan, S., Sorkin, R. D., & Itzkowitz, J. (submitted). A unified model of advice acquisition and use.

Macmillan, N. & Creelman, C. (1991). Detection Theory: A user's guide. Cambridge, England: Cambridge University Press.

Montgomery, D. A. & Sorkin, R. D. (1996). Observer sensitivity to element reliability in a multi-element visual display. Human Factors, 38, 484-494.

Morris, P. A. (1974). Decision analysis expert use. Management Science, 20, 1233-1241.

Morris, P. A. (1977). Combining expert judgments: A Bayesian approach. Management Science, 23, 679-693.

Murrell, G. W. (1977). Combination of evidence in perceptual judgment. In M. F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes in applied settings. (pp 169-201). New York: Academic.

Nitzan, S., & Paroush, J. (1982). Optimum decision rules in uncertain dichotomous choice situations. International Economic Review, 23(2), 289-297.

Papastravrou, J. & Athans, M. (1992). On optimal distributed decision architectures in a hypothesis testing environment. IEEE Transactions on Automatic Control, 37(8), 1154-1169.

Penrod, S., & Hastie, R. (1980). A computer simulation of jury decision making. Psychological Review, 87, 133-159.

Pete, A., Pattipati, K. R., & Kleinman, D.L. (1993). Tasks and organizations: A signal detection model of organizational decision making. Intelligent Systems in Accounting, Finance, and Management, 2, 289-303.

Pete, A., Pattipati, K. R., Kleinman, D. L., & Levchuk, Y. N. (1998). An overview of decision networks and organizations. IEEE Transactions on Systems, Man, and Cybernetics- Part C: Applications and Reviews 28(2), 173-193.

Robinson, D. E. & Sorkin, R. D. (1985). A contingent criterion model of computer assisted detection. In R. Eberts and C. G. Eberts (Eds.), Trends in Ergonomics/Human Factors. (Vol. II, pp 75-82). Amsterdam: North-Holland.

Shi, W., Sun, T. W., Wesel, R. D. (2001). Quasi-convexity and optimal binary fusion for distributed detection with identical sensors in generalized Gaussian noise. IEEE Transactions on Information Theory, 47, 446-450.

Sorkin, R. D. & Dai, H. (1994). Signal detection analysis of the ideal group. Organizational Behavior and Human Decision Processes, 60, 1-13.

Sorkin, R. D., Hays, C. J., & West, R. (2001). Signal detection analysis of group decision making. Psychological Review, 108, 183-203.

Sorkin, R. D., Luan, S., & Itzkowitz, J. (2004). Group decision and deliberation: A distributed detection process. In D. Koehler & N. Harvey (Eds.), Handbook of Judgment and Decision Making. New York: Oxford University Press.

Sorkin, R. D., West, R., & Robinson D. E. (1998). Group performance depends on the majority rule. Psychological Science, 9, 456-463.

Sniezek, J.A. & Henry, R. A. (1989). Accuracy and confidence in group judgment. Organizational Behavior and Human Decision Processes, 43(1), 1-28.

Stasser, G. & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. Journal of Personality and Social Psychology, 53, 81-93.

Stasser, G. & Vaughan, S. I. (1996). Models of participation during face-to-face unstructured discussion. In Witte, E. H. & Davis, J. H. (Eds.) Understanding Group Behavior. Hillsdale, NJ.: Lawrence Erlbaum Associates.

Swaszek, P. F. & Willett, P. (1995). Parley as an approach to distributed detection. IEEE Transactions on Aerospace and Electronic Systems, 31(1), 447-457.

Swets, J. A., Tanner, W. P., & Birdsall, T.G. (1958). Decision processes in perception. Psychological Review, 68, 301-340.

Tanner, W.P. & Sorkin, R.D. (1972). The theory of signal detectability. In Tobias, J.V. (Ed.), Foundations of Modern Auditory Theory, Vol II.. (pp. 63-98). New York: Academic Press.

Tanner, W. P., Jr. & Swets, J. A. (1954). A decision-making theory of visual detection. Psychological Review, 61, 401-409.

Viswanathan, R. & Varshney, P. K. (1997). Distributed detection with multiple sensors: Part I fundamentals. Proceedings of the IEEE, 85(1), 54-63.

Watson, W. E., Michaelsen, L. K. & Sharp, W. (1991). Member competence, group interaction, and group decision making: A longitudinal study. Journal of Applied Psychology, 76, 803-809.

Wickens, T. D. (1992). Elementary Signal Detection Theory. London: Oxford University Press.

Zarnoth, P. & Sniezek, J. A. (1997). The social influence of confidence in group decision making. Journal of Experimental Social Psychology, 33, 345-366.

BIOGRAPHICAL SKETCH

Jesse Itzkowitz obtained his undergraduate degree in psychology from the University of Florida in 2000. He remained at the University of Florida and continued his studies in cognitive psychology under the direction of Dr. Robert D. Sorkin. In 2003, he earned his master's degree with his initial work modeling deliberative groups. Jesse is expected to receive his doctoral degree from the University of Florida in August 2005. Afterward, Jesse will begin work towards a second doctoral degree in marketing at the Warrington College of Business Administration.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert D. Sorkin, Chairman
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ira Fischler, CoChairman
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark Fondacaro
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Joseph W. Alba
Distinguished Professor of Marketing

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 2005

Dean, Graduate School